Operations Research Letters 43 (2015) 461–466

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Network capacity control under a nonparametric demand choice model

M. Hosseinalifam [a,*], P. Marcotte [b], G. Savard [c]

[a] ExPretio Technologies, 200 Laurier Avenue West, Suite 400, Montréal, Québec, Canada H2T 2N8

[b] DIRO, Université de Montréal, C.P. 6128, succursale Centre-Ville, Montréal, Québec, Canada H3C 3J7

[c] MAGI, École Polytechnique de Montréal, C.P. 6079, succursale Centre-Ville, Montréal, Québec, Canada H3C 3A7

ARTICLE INFO

Article history:
Received 2 October 2014
Received in revised form
26 June 2015
Accepted 26 June 2015
Available online 3 July 2015

Keywords:
Revenue management
Customer choice behavior
Mixed integer programming

ABSTRACT

This paper addresses a dynamic resource allocation problem which has its roots in airline revenue management, and where customers select the available product that ranks highest on a preset list of preferences. The problem is formulated as a flexible mathematical program that can easily embed technical and practical constraints, as well as accommodate hybrid (parametric–nonparametric) choice models. We propose for its solution a column generation algorithm whose performance, both in terms of solution quality and processing time, is assessed against that of alternative approaches.

1. Introduction

Dynamic resource allocation is the main element of revenue management (RM), the discipline whose aim is to improve a firm's profitability through efficient pricing and asset management. These issues involve the design of decision rules that, over the booking horizon, allow or deny access to products that use common resources, based on assumptions concerning the behavior of customers facing distinct options [9].

While traditional RM models assume cross-product independence, as well as independence from both capacity control strategies and from the state of the market, more sophisticated discrete choice models have become increasingly popular [9]. These posit that customer behavior is dictated not only by product availability, but also by products' attributes (price, quality, restrictions, willingness to pay, etc.) [10]. In turn, alternative parametric models that obviate some of these models' limitations have been proposed [5].

In this paper, departing from the parametric approach, we consider a framework where the customer population is partitioned into segments, each segment being associated with an ordered list of product preferences that includes the 'no purchase' option. This demand model is then embedded within a capacity control system. More precisely, given a stochastic arrival process that governs each segment, acceptance rules that determine the optimal set of products offered in each time period are obtained through the solution of a deterministic mathematical program. An important feature of the model is its flexibility with respect to additional constraints. In particular, it can accommodate arbitrary topologies that go beyond the traditional hub-and-spoke architecture, as well as user-specific constraints.

Our contribution is twofold. First, we propose a nonparametric model for choice-based revenue maximization, through the specification of the optimal sets of products that are made available at each booking period. Next, in the view that the number of variables grows exponentially with the number of products, we develop an efficient column generation algorithm that exploits the specific structure of the choice model, and has the capability of addressing real-life instances.

The structure of the paper is as follows. In Section 2, we review the main concepts of choice-based demand models, contrasting the parametric and nonparametric approaches. In Section 3, we introduce our mathematical programming framework. In Section 4, we develop a column generation scheme for its solution; in particular, we provide an efficient algorithm for tackling the nonconvex subproblems. In Section 5, we illustrate through computational experiments that our approach can address realistic instances, and provide a comparison with alternative approaches from the RM literature. Finally, in the concluding section, we outline the challenges that remain to be addressed.

* Corresponding author.
E-mail addresses: Morad.Hosseinalifam@expretio.com (M. Hosseinalifam), Marcotte@iro.umontreal.ca (P. Marcotte), Gilles.Savard@polymtl.ca (G. Savard).

http://dx.doi.org/10.1016/j.orl.2015.06.012
0167-6377/

2. Choice modeling

In this section, we briefly survey choice modeling. Indeed, a key issue in network revenue management is that of estimating the probability $P_j(S)$ that a product j be selected by an arriving customer, given that a set S of products is on offer. Two main classes of models have been proposed for its solution. Parametric choice models are built upon the Random Utility Maximization paradigm [1], whereby products are assigned attributes, and customers select the product that maximizes their own utility, expressed as a weighted sum of the attributes' values. Depending on the statistical model underlying the selection process, one derives a variety of models: multinomial logit (MNL), nested logit, mixed logit, probit, generalized extreme value, etc.

Although random utility models are easy to understand, embed detailed information about products' features, and allow the accurate estimation of utilities, they yet suffer serious flaws. First, the choice of an appropriate parametric structure may not be obvious and, once a structure is adopted, the model is not flexible with respect to perturbations in the available information [5]. Next, specific random utility models have specific drawbacks. For example, MNL's independence of irrelevant alternatives property yields unrealistic substitution patterns, while the more sophisticated nested or mixed logit models are computationally challenging, both from an estimation and assignment viewpoint.

In contrast, nonparametric choice models are driven by historical data and do not assume specific probability distributions. They are highly flexible, dynamic, and provide more precise estimates of customer's choice behavior [5]. Due to the availability and increasing accuracy of large amounts of historical data, these choice models have been gaining in popularity and interest.

In the nonparametric choice model adopted in this paper, we assume that each demand segment is characterized by an Ordered Preference List (OPL), whereby customers select the available product that ranks highest on their OPL, possibly leaving the market if no available product belongs to the list [3,4]. The concept of OPL was first introduced in [8], while [5,12] proposed different procedures to estimate a non-parametric choice model. Within this framework, [11,3] used the concept of OPL to formulate a choice based capacity control model, and proposed for its numerical solution a stochastic gradient algorithm; [6] developed an algorithm to compute optimal assortments under a nonparametric choice model of demand; [8] proposed a model to compute optimal retail assortments, in an environment where customers adapt dynamically to available stocks; [4] have proposed OPL-based linear stochastic formulations of the revenue maximization problem which, unfortunately, become intractable as the number of scenarios grows. To tackle this problem, they proposed a heuristic approach and estimated a linear approximation of their stochastic model. This method reduces the processing time, however, and reduces the quality of the solution. Our approach shares several features with this work, while lifting its computational limitations.

3. Problem formulation

Let us consider a system where an arrival stream of customers follows a Poisson process with rate λ. Whenever a customer shows up, he selects the available product that ranks highest on his preference list. The aim of the model is to determine, over a finite planning horizon, the set of products to be offered at any given 'booking' period, in order to maximize total revenue. Of course, a product can only enter the offer set if the amount of resources required does not exceed the residual amount available.

The main parameters underlying the dynamic RM model are the followings:

- T: ordered set of time (booking) periods indexed forward
- J: set of products
- r_j: revenue associated with product $j \in J$
- I: set of resources
- c_i: initial amount of resource $i \in I$
- L: set of customer segments
- p_l: proportion of customers belonging to segment l
- $\lambda_l = \lambda p_l$: arrival rate of segment $l \in L$
- $P_l = \exp(-\lambda p_l)$: probability of an arrival issued from segment $l \in L$ within an arbitrary time period
- $O^l = \{j^l_1, j^l_2, \ldots, j^l_{K_l}\}$: ordered preference list (OPL) of products associated with customer segment $l \in L$
- a_{ij}: Boolean constant indicating whether resource i is used by product j ($a_{ij} = 1$) or not ($a_{ij} = 0$). The matrix A whose elements are the a_{ij}'s is referred to as the product-resource incidence matrix or, simply, the incidence matrix.
- $S \in 2^J$: set of products, possibly including the 'null' product
- $O^l(S) = \{j^l_1(S), j^l_2(S), \ldots, j^l_{K_l(S)}(S)\} \subseteq O^l$: ordered preference list (OPL) of cardinality $K_l(S)$ associated with offer set S and customer segment $l \in L$.

Assuming that there is at most one arrival within any given time period, the probability of choosing product j when set S is offered is equal to

$$P_j(S) = \sum_{l:j^l_1(S)=j} P_l, \tag{1}$$

where $j^l_1(S)$ is the first available preferred product of segment l among those belonging to the offer set S. This yields the expected revenue

$$R(S) = \sum_{j\in S} P_j(S) r_j \tag{2}$$

and the expected capacity usage of set S

$$Q_i(S) = \sum_{j\in S} P_j(S) a_{ij}. \tag{3}$$

The variables of the model are the indicators $X_t(S)$, which specify whether the subset of products S is offered or not in period t. For the sake of computational tractability, we allow these binary variables to assume fractional values. Based upon these definitions, letting $X = (X_t(S))_{t,S}$, and denoting by $|E|$ the cardinality of a generic set E, the model can be expressed as the linear program

$$\text{LP:} \qquad \max_X \lambda \sum_{t\in T} \sum_{S\in 2^J} R(S) X_t(S)$$

subject to

$$\sum_{t\in T} \sum_{S\in 2^J} \lambda Q_i(S) X_t(S) \le c_i \quad \forall i \in I, \tag{4}$$

$$\sum_{S\in 2^J} X_t(S) \le 1, \quad \forall t \in T, \tag{5}$$

$$0 \le X_t(S) \le 1 \quad \forall t \in T,\ S \in 2^J, \tag{6}$$

where the number of decision variables $2^{|J|} - 1$ is exponential (the empty set is excluded).

Note that the above program is similar to the customer MNL-based deterministic linear programming model (CDLP) considered by [7,2], the main differences being the way we model customer's choice behavior, and the way we compute the related probabilities that lead to the values of $R(S)$ and $Q_i(S)$. Note also that, in contrast with [7,2], our decision variables are related to individual time periods, thus allowing a finer control over the individual booking periods. Finally, the use of ordered preference lists allows